Perceptually Based Theory for World Music Tunings

Jay Rahn, York University

Annual Conference, Society for Music Theory,
Minneapolis, 2011.
http://societymusictheory.org/events/meeting2011/igactivities

The larger research program, of which the present report is a part, involves approaching the analysis of world music from the vantage point of perception. With regard to tuning and scales, there have been three main approaches: via abstract numbers, via empirical measurements, and via perceptual responses.

Since the pioneering work of Alexander Ellis in the 1880s, comparative musicologists, ethnomusicologists, acousticians, and music psychologists have measured fundamental frequencies and other partials in sonic spectra to characterize tunings empirically. Since at least as early as ca. 425 BCE, music theorists have employed abstract numbers to prescribe and describe tunings and scales. And since at least as early as ca. 1850 BCE, tunings have been prescribed and described in terms of perceptual responses.

Whereas each of these three main approaches can be helpful in making sense of music, they are quite distinct. Nonetheless, these three distinct approaches have often been conflated and an aim of the present study is to disentangle them with a view to analyzing music perceptually.

The present report illustrates this approach by focusing on equipentatonic tunings in general and Central Javanese *sléndro* in particular. Early transcriptions and descriptions of music in *sléndro* tuning, by Raffles, Crawford, and Figée (as reported by Ellis), employed European-derived notation and letter names that correspond to the solmization syllables doh, re, mi, sol, and la. Such tones correspond to numerical formulations in recent music theory: e.g., interval-class set 5-35 in Forte's listing, the 'usual pentatonic' in Clough and Douthett's study of maximally even sets, and in Carey and Clampitt's formulation, an instance of '***non-degenerate*** well-formed scales.'

In contrast, glossaries in the *Garland Encyclopedia of World Music* employ the term 'equipentatonic' and specify, in a formulation that can be traced to measurements made by Ellis, that 'equipentatonic' refers to a 'pitch inventory with five ***equally spaced*** pitches to the octave.' In Carey and Clampitt's numerical formulation, such a pitch inventory would be an instance of a '***degenerate*** well-formed scale.'

Whereas glossaries in the *Garland Encyclopedia* state baldly that equipentatonic pitches are equally spaced, other passages in this authoritative reference work, as well as corresponding portions of the similarly authoritative *Grove Music Online and Oxford Music Online*, have followed the lead of Klaus Wachsmann's term 'pen-equidistant' (1967) and qualify the notion of equality by means of such phrases as '**nearly equidistant**,' '**nearly equal**,' '**almost equal**,' and '**roughly equal**.' Indeed, *Grove Music Online*'s entry for 'Interval' is even more guarded, saying merely that the *sléndro* scale of the Indonesian gamelan '**sometimes approximates**' to a division of the octave into five equal parts.

Not only have there been three distinct ways of construing such tunings as *sléndro*. As well, Albrecht Schneider has made the strong claim that 'many of the tone measurements carried out on dozens of gamelan instruments … are ***at odds with*** the physical and psychophysical nature of the sounds gamelan instruments such as gong chimes and other metallophones actually produce.' By way of explanation, he goes on to say that, 'due to inharmonic spectra and non-periodic time functions characteristic of the sounds of idiophones,' the practice of 'equating the pitch of a tone with a single frequency … ***will not*** work very well, and may lead to interpretations which seem, at best, problematic if not obsolete.'

I have compared Schneider's claims with tones one can hear and measure at a readily accessible website that documents 'Kyai Parijata,' a two-century old Central Javanese gamelan that has been housed and played for more than 40 years in Delft. As individual tones of this gamelan differ considerably in amplitude, the leftmost columns of ***Figure 1*** display the ranges of fundamental frequencies for louder and softer tones that have been measured by means of software employed by the gamelan's leader. The rightmost columns of ***Figure 1*** display the frequencies and amplitudes of the illustrative tones for which the website provides wav files, and which I measured by means of widely available inexpensive software, namely, Transcribe!

Instruments:	**Measurements by van Oldenborgh**:		**Lowest Peaks in Transcribe! Software**	
	minima in Hz	**maxima in Hz**	**in Hz**	**dB**
Saron	1084.88	1084.91	1084.91	-29.3
Barung	938.41	938.42	938.47	-21.5
	820.12	820.34	819.31	-22.4
	710.38	710.44	710.37	-21.2
	614.79	615.35	614.50	-21.9
	535.21	535.46	534.02	-15.8
	462.72	462.89	463.01	-16.3
Saron	535.66	535.70	536.48	-6.9
Demung	463.24	463.27	463.01	-9.6
	402.80	402.81	403.30	-10.2
	351.56	351.68	351.28	-10.8
	302.28	302.37	301.78	-15.7
	264.10	264.24	264.08	-10.7
	230.55	230.68	230.55	-5.5
Slenthem	264.53	264.54	264.08	-9.2
	231.08	231.09	231.08	-14.5
	199.56	199.57	199.43	-14.1
	174.70	174.74	174.51	-13.5
	151.06	151.09	151.31	-14.6
	129.42	129.43	129.39	-14.9
	115.02	115.02	115.33	-14.9

Figure 1.a. Comparison of fundamental frequencies slenthem, saron demung, and saron barung keys of gamelan Kyai Parijata of Delft, as measured by Geert Jan van Oldenborgh and by means of Transcribe! software.

For the tones measured by van Oldenborgh, 'The frequency was measured by taking a 2^16 sample FFT (1.4s) starting at the beginning of the sound. This FFT was fitted over 10 bins by a Bessel function with undetermined amplitude, frequency and width, plus constant background,' and for repeated measurements, the error is 'normally around 0.1Hz,' i.e., < 0.2 cents. (http://www.marsudiraras.org/gamelan/Delft.notes)

For the lowest peak frequencies measured by means of Transcribe! software, the measurement gradations for *slenthem*, *saron demung*, and *saron barung* tones are uniformly ± 0.46%, i.e., ± 8 cents.

The accompanying ***audio example*** comprises the *saron barung*, *saron demung*, and *slenthem tones*, in that order, from highest to lowest: http://yorkspace.library.yorku.ca/xmlui/handle/10315/28322

Instruments:	Measurements by van Oldenborgh:		Lowest Peaks in Transcribe! Software:	
	minima in Hz	maxima in Hz	in Hz	dB
Bonang Panerus	2492.30	2498.90	n.a.*	-6.3
	2130.30	2130.70	n.a.*	-10.2
	606.77	1862.90	1858.87	-12.7
	1621.30	1626.80	1626.60	-16.6
	1422.50	1423.80	1423.86	-12.5
	1240.00	1246.00	1245.51	-16.6
	1086.60	1087.60	1084.88	-15.2
	939.95	941.00	940.63	-14.9
	815.80	819.16	819.31	-21.7
	709.77	710.29	707.11	-10.6
	615.79	620.30	617.33	-8.1
	534.06	536.23	536.48	-5.8
Bonang Barung	1239.70	1240.60	1239.79	-8.8
	1083.99	1084.90	1084.88	-19.3
	935.23	937.43	936.31	-25.0
	820.52	820.96	821.20	-23.6
	702.47	706.70	710.37	-18.3
	610.72	624.01	618.76	-5.4
	535.46	535.85	536.48	-6.7
	460.62	470.23	469.45	-8.9
	403.39	409.71	408.90	-1.8
	352.27	352.30	352.90	-7.3
	300.53	300.90	300.40	-11.3
	264.26	264.89	264.08	-6.7

Figure 1.b. Comparison of fundamental frequencies of *bonang barung* and *bonang panerus* keys of gamelan Kyai Parijata of Delft, as measured by Geert Jan van Oldenborgh and by means of Transcribe! software.

Concerning van Oldenborgh's frequency measurements, see Figure 2.a (above). For the lowest peak frequencies measured by means of Transcribe! software, the measurement gradations for *bonang barung* and *bonang panerus* tones are uniformly ± 0.46%, i.e., ± 8 cents. Note also (*, above) that Transcribe! software measures frequencies only up to ca. 2000 Hz.

Also note the anomalous measurement of the third highest bonang panerus tone (***highlighted*** above).

The accompanying ***audio example*** comprises the boning panerus and bonang barung tones, in that order, from highest to lowest: http://yorkspace.library.yorku.ca/xmlui/handle/10315/28321

As ***Figure 1*** shows, there is very little difference between the frequencies measured by both methods. In order to compare these measurements with what one hears, I have employed Audacity freeware to generate sine tones having the same fundamental frequencies as those that resulted from the Transcribe! measurements and juxtaposed the website's illustrative sound files with the respective sine tones.

In the audio examples that accompany ***Figure 1***, each gamelan tone is immediately followed by its respective sine tone. Unless my ears deceive me, each sine tone is heard as matching quite closely in pitch the immediately preceding gamelan tone. To be sure, within the gamelan tones one can also hear partials higher than the fundamental: as Schneider says, some occur during the onset and others are audible during the gamelan tones' relatively 'steady state.' However, within each gamelan tone, there is a most salient partial that corresponds in pitch to the sine tone.

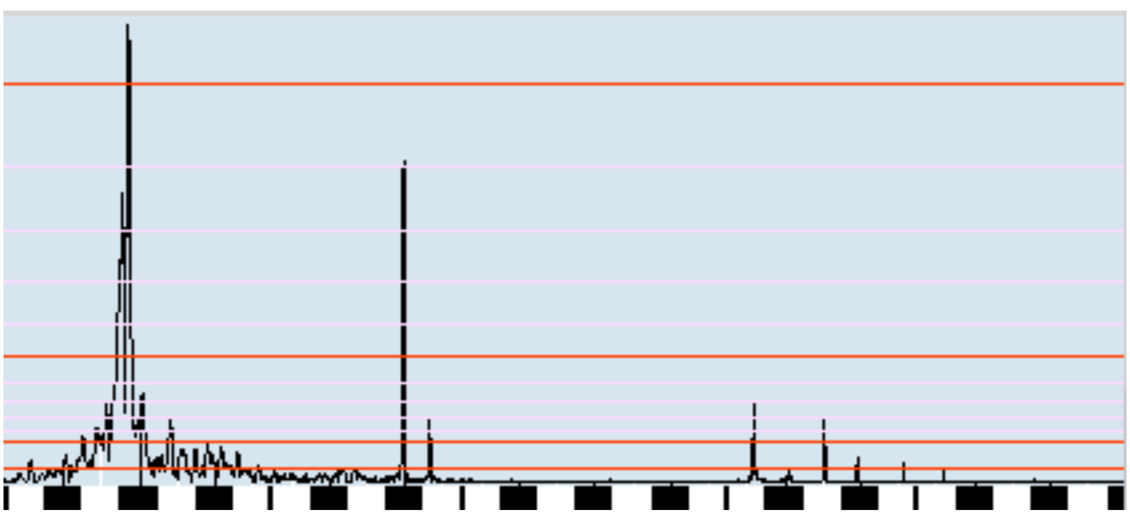

Figure 2. Graphic display produced by Transcribe! Software for Frequencies (x-axis) and their Amplitudes (y-axis) in the Fourth Lowest *Bonang Barung* tone of Figure 1(b). The keyboard portion of the graphic ranges from F4 to B-flat6 (i.e., ~349-2000 Hz).

Significantly, according to the graphic displays provided by Transcribe! software, the partial of each tone that matches in pitch the sine tone corresponds to the lowest peak frequency and this lowest peak frequency is also the peak frequency of greatest amplitude E.g., in the Transcribe! graphic display of ***Figure 2***, a frequency near G#4 corresponds to the fourth lowest of the *bonang barung* tones in ***Figure 1.b*** and clearly has the greatest amplitude. This observation validates studies that, since the 1930s, have employed stroboscopic tuners to measure fundamental frequencies, for stroboscopic tuners are best at specifying the peak frequency of greatest amplitude for an individual tone and the perceptual comparisons just demonstrated show that what one would consider 'the pitch' of an individual tone on the basis of its lowest peak frequency of greatest amplitude corresponds quite closely to 'the pitch' of a sine tone having the same frequency.

Conversely, this comparison also validates studies that, since the 1880s, have employed tuning forks or well calibrated monochords to measure fundamental frequencies. For the latter involve comparing an individual tone with a tone produced by a tuning fork or a monochord and determining by perception whether the two tones match in pitch. Whether an electronically generated sine tone, or a tone produced by a tuning fork or a monochord is the perceptual comparator, its fundamental frequency serves as a ***surrogate*** for what one terms 'the pitch' of a tone. Here one should emphasize that apart from the highly problematic case of so-called 'absolute pitch,' one cannot speak properly of 'the pitch' of a tone. More properly, two tones—or more generally two parts of tones—are heard as matching, or not matching, in pitch.

In general, if a sine tone and the lowest peak frequency of such a gamelan tone have the same fundamental frequency they are heard as matching in pitch. That is, having the same fundamental frequency is a sufficient condition for their being heard as the same in pitch. However, having the same fundamental frequency is not a necessary condition. Otherwise, there would be no reason to specify the acoustically measured conditions for a so-called '***just noticeable difference***' ***(jnd)*** in frequency. Moreover, as considerable research in psychology has shown, jnd's for pitch, as for other perceptual variables, are a matter of ***individual differences***, which can be considerable.

Similarly, if the fundamental frequencies of two such tones constitute the same ratio as the fundamental frequencies of another pair of such tones, the two pairs of tones are generally heard as matching ***pitch-intervallically***. As with pitch, constituting the same fundamental-frequency ratio is a sufficient, but not a necessary, condition for hearing two tone-pairs as matching pitch-intervallically. Are *sléndro* tones, then, ‘equally spaced’?

Figure 3 displays the ranges of values of the fundamental-frequency ratios formed by the pairs of tones in Kyai Parijata that span 0, 1, 2, 3, …, 20 steps, i.e., successive keys and pots. In Figure 3, one can see that, as calculated in cents, i.e., hundredths of an equally tempered semitone, the fundamental-frequency ratio of the largest so-called ‘***unison***,’ i.e., the largest 0-step interval (or ‘prime’), is 184 cents smaller than the fundamental-frequency ratio of the smallest so-called 2nd, i.e., the smallest 1-step interval, and similarly for the largest 1-step interval and the smallest 2-step interval, and so forth. Among all the intervals, i.e., from 0 to 20 steps, the more steps an interval spans, the larger ***all*** of its instances are.

number of steps:

	0	**1**	**2**	**3**	**4**	***5***	**6**	**7**	**8**	**9**	***10***
min	[0]	199	454	693	932	***1179***	1414	1657	1904	2151	***2382***
max	24	271	518	749	1000	***1235***	1474	1729	1968	2215	***2454***
difference	175	183	175	183	179	179	183	175	183	167	

	(10)	**11**	**12**	**13**	**14**	***15***	**16**	**17**	**18**	**19**	***20***
min	(2382)	2622	2877	3139	3378	***3610***	3865	4096	4343	4582	***4813***
max	(2454)	2697	2940	3195	3426	***3681***	3912	4152	4382	4614	***4813***
difference	168	180	199	183	184	184	184	191	200	199	

Figure 3. Minimum (min) and maximum (max) fundamental-frequency ratios, in cents, of pairs of tones comprising 0 to 20 steps in Kyai Parijata.
Cents measurements are based on the fundamental frequencies measured by Transcribe! software.

The lowest rows specify the difference between the maximum values of intervals that span n = 0, 1, 2, …, 19 steps and the respective minimum values of intervals spanning n+1 = 1, 2, 3, …, 20 steps. (The smallest difference is 167 cents.)

Values for the 5-step ('octave') intervals of and their supplementary intervals of 10, 15, and 20 steps are ***highlighted***.

To employ David Rothenberg's and Norman Carey's terms, Kyai Parijata's *sléndro* tuning is, respectively, 'strictly proper' or '***generically ordered***.' Indeed, if one relativizes the notion of generic ordering, as Carey's study of scale candidacy does implicitly, Kyai Parijata's *sléndro* tuning can be considered '***mensurably generically ordered***.' With regard to measurement, the steps in Kyai Parijata's *sléndro* tuning are definitely not 'equally spaced.' However, an assessment of generic ordering bypasses problematic issues of equality with regard to measurement and substitutes the notion that, as measured, Kyai Parijata's *sléndro* tuning is generically ordered.

As mentioned above, an equally spaced tuning is a degenerate well-formed tuning. In such a tuning there are 5 classes of intervals. One of the classes spans 0 steps and its specific magnitude is 0, another spans 1 step and its specific magnitude is 1/5 the magnitude of the modulus, i.e., 1/5 of the so-called 'octave,' and so on. Such a tuning is highly unified insofar as the number of steps an interval spans varies directly, in a one-to-one manner, with its specific magnitude. However, such a tuning's degenerate well-formed structure is purely numerical, i.e., ***abstract***, and, in principle, ***no measurement can verify it***.

Kyai Parijata's *sléndro* tuning is highly unified in another way: For each number of steps there is a group of intervals that are 'the same' in that all intervals that span that number of steps are smaller in specific magnitude than all the intervals that span more steps. Moreover, for such a mensurably generically ordered tuning, specific magnitudes are assessed by measurement. Further, whereas measured sameness is in general, a sufficient, but not a necessary, condition for perceptual sameness, measured difference is a necessary, but not a sufficient condition for perceptual difference. However, as ***Figure 3*** shows, to deny that Kyai Parijata's *sléndro* tuning is not only mensurably generically ordered but also ***perceptibly generically ordered*** would be to claim that persons do not hear pairs of tones whose measured fundamental-frequency ratios differ by at least 167 cents as differing pitch-intervallically.

For purposes of musical analysis and relative to such a 167-cent difference, all intervals in Kyai Parijata that span a particular number of steps can be considered to instances of a grouping that is unified by analogical relationships between numbers of steps and auditory magnitudes. As well, there is room in an analysis for the possibility that any interval that spans a particular number of steps is heard as larger than one or more intervals that span the same number of steps—just as in the European-derived distinction between a C-sharp and a 'sharp C' or between a minor 3rd and a small major 3rd.

If no such difference is heard, all intervals that span a particular number of steps are heard as the same and Kyai Parijata's tuning can be regarded as ***degenerate well-formed***—albeit in a ***non-numerical*** sense of perceived ***same-as*** relationships, or more precisely, ***perceptible analogical relationships***, rather than abstract numbers.

If, however, such differences are heard, they can be considered ***nuances*** within analogical groups of intervals spanning the same number of degrees. To judge from previous studies of tuning perception, such differences or their absence would be a matter of ***differences among individual persons***. And as ***Figure 3*** shows, one would anticipate such differences in individual responses to Kyai Parijata's tones to correspond to differences of less than 167 cents.

Finally, if such differences occur among corresponding steps in all registers, they might be considered instances of ***sub-generic*** ordering, as in the case of the usual pentatonic, which is generically ordered and where, in each register, 3 of the 1-step intervals are heard as smaller than the other 2, 1 of the 2-step intervals is heard as smaller than the other 4, and so forth.

References Cited

Carey, Norman. 1998. *Distribution modulo 1 and musical scales*. Ph.D. dissertation, University of Rochester: https://www.academia.edu/3385414/Distribution_modulo_one_and_musical_scales

______ and David Clampitt. 1989. Aspects of Well-Formed Scales. *Music Theory Spectrum* 11: 187–206.

Clough, John and Jack Douthett. 1991. Maximally Even Sets. *Journal of Music Theory* 35: 93-173.

Ellis, Alexander J. 1884-85. On the Musical Scales of Various Nations. *Journal of the Royal Society* 33: 485-527.

______ and Alfred J. Hipkins. 1884. Tonometrical Observations on Some Existing Non-Harmonic Musical Scales. *Proceedings of the Royal Society of London* 37: 368-385.

Lindley, Mark, Murray Campbell, and Clive Greated. 2001. Interval. *Grove Music Online.* Oxford, UK: Oxford University Press. (subscription access)

van Oldenborgh, Geert Jan van. n.d. Gamelan Group Marsudi Raras. http://www.marsudiraras.org/

Rothenberg, David. 1978. A Model for Pattern Perception with Musical Applications Part I: Pitch Structures as Order-Preserving maps. *Mathematical Systems Theory* 11: 199–234.

Schneider, Albrecht. 2001. Sound, Pitch, and Scale: From 'Tone Measurements' to Sonological Analysis in Ethnomusicology. *Ethnomusicology* 45/3: 489-519.

Seventh String Software. Transcribe!: http://www.seventhstring.com/

Stone, Ruth M., ed. 1994. *The World's Music: General Perspectives and Reference Tools*. Vol. 10 of *The Garland Encyclopedia of World Music Online*. Alexandria, Virginia: Alexander Street Press. (subscription access)

Wachsmann, Klaus. 1967. Pen-Equidistance and Accurate Pitch: A Problem from the Source of the Nile. in Ludwig Finscher and Christoph-Hellmut Mahling, eds. *Festschrift für Walter Wiora zum 30. Dezember 1966.* Kassel: Bärenreiter 583–592.